7/17/03

In the Specification:

Please substitute the following paragraph(s) of the specification for their corresponding pending paragraph(s) under 37 C.F.R. 1.121(b)(1):

1. Pending *Cross Reference To Related Applications* of the specification:

A1

Priority is claimed from U.S. Provisional Application No. 60/311,857, filed August 13, 2001. Cross-reference is made to U.S. Patent Application Serial No. 09/543,962, entitled "Meta-Document And Method Of Managing", and U.S. Patent Application Serial No. 09/928,619 entitled "Fuzzy Text Categorizer", which are both hereby incorporated herein by reference. In addition, cross-reference is made to the following U.S. Patent Applications that (a) are concurrently filed with this patent application, (b) are assigned to the same assignee as this patent application, (c) are incorporated in this patent application by reference, and (d) claim priority to U.S. Provisional Patent Application Serial No. 60/311,857, filed August 13, 2001: U.S. Patent Application Serial No. 09/683,238, entitled "Meta-Document Management System With Personality Identifiers"; U.S. Patent Application Serial No. 09/683,239, entitled "Meta-Document Management System With Document Identifiers"; U.S. Patent Application Serial No. 09/683,240, entitled "Meta-Document Management System With Transit Triggered Enrichment"; U.S. Patent Application Serial No. 09/683,235, entitled "System For Automatically Generating Queries"; U.S. Patent Application Serial No. 09/683,241, entitled "System For Propagating Enrichment Between Documents"; U.S. Patent Application Serial No. 09/683,242, entitled "Document-Centric System With Auto-Completion And Auto-Correction"; U.S. Patent Application Serial No. 09/683,236, entitled "Meta-Document Management System With User Definable Personalities".

IWO IWO IWO IWO IWO IWO IWO 09/22/03

2. Pending paragraph number **442**:

A2

This section describes a mechanism that uses an information space surrounding a document to provide an improved (e.g., more accurate and more stylish) document-centric auto-completion system and auto-correction system that can be used

A2

during content creation. Document auto-completion saves a user from having to retype text (and other document content such as graphics) and related markup such as hyperlinks, bibliographic entries etc., by providing suggestions of words that have been used previously in a contextually similar manner. Document auto-correction provides a textual correction system that dynamically updates the information space as corrections are made or accepted.

3. Pending paragraph number **460**:

A3

Subsequently at 4506, the module 4406 waits for a signal from text editor 4314 that document content 4203 has been added and/or edited. At 4508, the information space is updated based on the added and/or edited document content. At 4510, the updated information space (i.e., added and/or edited document content and enrichment associated therewith) is processed for entities that could potentially be used for auto-completion. At 4512, if extracted entities are deemed to be appropriate for auto-completion, then they are indexed and inserted into the database of entities 4214; otherwise, or upon completion of 4512, the service 4406 waits for additional signals from the editor 4314.

4. Pending paragraph number **467**:

A4

Also, other factors such as the length of entities, highlighting information (i.e. are headings, bold, hyperlinked, etc.), markup information (such as hyperlinks, footnotes etc.), location of the entity in a document, its frequency in a document (or within a corpus) could be used in any combination to determine the utility of inserting the entity into the entity completion database. Those entities with a utility above a certain threshold are selected and inserted into the entity database. In one embodiment, the utility of an entity is determined using a weighted linear combination of factors as set forth below:

A4

$$Utility(entity) = \sum_{factors} weight_{factor}(factor) =$$
$$weight_{bold}(bolded(true = 1; false = 0)) +$$
$$weight_{italic}(italic(true = 1; false = 0)) +$$
$$\vdots$$
$$weight_{uppercase}(upercase(true = 1; false = 0)) +$$
$$weight_{location}\left(1 - \frac{location\ of\ word}{document\ length}\right) +$$
$$weight_{frequency}\left(\frac{frequency\ of\ word\ occuring\ in\ document}{hightest\ frequency\ of\ any\ word\ in\ document}\right) +$$
$$weight_{corpus}\left(\frac{frequency\ of\ word\ occuring\ in\ corpus}{hightest\ frequency\ of\ any\ word\ in\ corpus}\right).$$

5. Pending paragraph number **472**:

A5

Subsequently, a query is formulated at 4606 using the extracted context information and string fragment. In one embodiment, the query can simply be the string fragment. In alternative embodiment, the query can be expanded using various contextual information that may lead to more accurate suggestions for completion. For example, the auto-completion system could process the sentence of which the string fragment is a member using linguistic processing tools such as XeLDA (Xerox Linguistic Development Architecture) described in U.S. Patent No. 6,321,372, which is incorporated herein by reference.

6. Pending paragraph number **475**:

A6

At 4608, the formulated query is submitted to the information retrieval system 4308 in the auto-completion module 4302. In operation, the information retrieval system 4308 locates matches subject to the constraints specified in the query using known matching techniques. The matched items are retrieved and ranked based on their level of appropriateness for completion (i.e., how well they satisfy the query constraints and possibly additional constraints such how near each matched item is to a previously completed item) at 4610. The top ranked match that contains the same (or similar)